Journal of Choice Modelling, 4(2), pp 1-19
www.jocm.org.uk

Properties of Internet and Telephone Data Collection Methods in a Stated Choice Value of Time Study Context

Maria Börjesson[1,*] ¶ Staffan Algers[2,†]

[1]Centre for Transport Studies
Royal Institute of Technology, Sweden

[2] Centre for Transport Studies
Royal Institute of Technology, Sweden

Received 28 January 2010, revised version received 2 October 2010, accepted 14 March 2011

Abstract

We analyse Internet and telephone Stated Choice (SC) survey methods in the context of the Swedish value of time study 2008. In this study, extensive piloting and follow-up surveys were undertaken to assure high quality data. We use these data and data from the main survey to analyse properties of the different data collection methods. One conclusion is that the Internet gives less random error in the SC data. On the other hand, the response rate drops when the Internet is the only response and recruiting mode. A mixed mode survey, where the Internet is the primary method but where respondents are knowingly subject to a telephone follow up survey, is found to give substantially higher Internet response rates. If the telephone follow-up does not include SC questions, the value of time result will still be biased. However, a large part of this bias seems to be explained by socio-economic data, such as income and age, which are cheaper to collect.

Keywords: Stated Choice, data collection, value of time, mixed mode surveys, response error

* Corresponding author, T: +46 8 7906837, F: +46 8 7907002, maria.borjesson@abe.kth.se
† T: +46 8 7906074, F: +46 8 7907002, algers@infra.kth.se

ISSN 1755-5345

1 Introduction

Using Internet-based Stated Choice (SC) surveys has now become increasingly popular as a tool for elicitation of monetary valuations of time savings and other attributes. Interviewing over the Internet has become considerably cheaper that interviewing by telephone. The Internet has the further advantage of making it possible to customize SC surveys, as compared to mail-back questionnaires. Mixed mode surveys, in which both the Internet and an alternative interview mode are used, have been increasingly used to increase response rates. Doing so invokes the question of how the use of different modes affects the results. This paper aims at contributing to this literature by exploring response bias and quality of SC responses in Internet surveys as compared to telephone-based surveys and to a combination.

We do this in the context of the 2008 Swedish Value of Time study. In this study, pilots were undertaken to compare telephone and Internet surveys. In telephone interviews, the respondents were asked to write the travel cost and travel time of the stated choices read out by the telephone interviewer on supplied paper sheets, to visually see the alternatives as compared to the Internet where the alternatives are directly visible on the screen.

The main study was carried out as a mixed mode survey. The pilots and main studies enable us to compare different forms of survey administration methods and assess how they can effectively be combined. The analysis concentrates on the following issues:

1. Does the error in responses to SC questions answered by telephone and by Internet differ? It has been found in other studies, that time pressure caused by the interview situation means that respondents do not take the time to fill in the SC sheet in order to see both alternatives, and therefore they do not consider the alternative enough. This may cause a higher degree of response error in telephone surveys.
2. How are response rates affected when using Internet rather than telephone surveys?
3. Do respondents who are not able to respond via Internet differ from Internet respondents in terms of socio-economic characteristics or VTT?

Ultimately, we try to answer the relevant question if the extra effort to perform full scale follow-up interviews by telephone in Internet surveys is worth the costs.

Data collection methods have been changing rapidly since the penetration of Internet and mobile phones is increasing, while the number of fixed telephone subscriptions is declining. In Sweden, about 71 percent of all households had a landline telephone in 2007. In addition, about 400,000 of all households (about 8 percent) were using IP-based telephony. A very large proportion of individuals had a mobile telephone. The number of mobile telephone subscriptions was more than 10 million in a nine million population. Hence, almost the entire population can be

reached by telephone. All the same, telephone numbers cannot be matched to a random population sample for approximately 30 percent, because there is no joint telephone directory including all mobile and IP-based telephone subscriptions, and because the fixed telephone subscriptions are sometimes only in the name of another household member. Hence, in telephone-based interviews, about 30 percent of respondents are bound to be excluded from the samples because their telephone number is not registered.

Of the population in the age group 18-74 years, 84 percent had access to Internet at home in 2007. Of those in employment, 72 percent had access to the Internet at their work place. Almost all students have Internet access from school. Hence, we expect about 16 percent of the population to drop out of Internet surveys because of lack of Internet access at home. However, people could have access to the Internet through friends, workplaces, libraries, or mobile phones.

The paper is organized as follows. Section 2 includes a brief literature review. Section 3 outlines the method and Section 4 summarizes the data. Estimation results are presented in section 5, while section 6 concludes the paper.

2 Literature

Data collection methods have been subject to much research, to an increasing extent also related to Internet-based methods. Some of this research is also related to discrete choice surveys. For example, Sethuramana et al. (2005) compare mail-back surveys to Internet-based surveys with respect to preference elicitation for attributes related to wireless telephone handsets. They observe significant preference differences between mail-back and online data collection methods. The online method was found to be superior to the mail-back method in terms of internal consistency and predictive validity. Olsen (2009) also compares Internet and mail-back surveys, in this case for eliciting preferences for protecting different types of landscape from road encroachment when building new motorways in Denmark. The performances of the two samples (148 persons using Internet and 152 using mail back) are compared over six different criteria: response rates, protest responses, demographics, preferences and WTP, estimation precision, and certainty in choice. No significant differences in WTP are found, although Olsen suggests that analysts should be aware that choosing Internet over mail could be accompanied by a survey mode effect because of differences related to some of the studied criteria.

Nossum (2004) gives an overview of different SC studies related to security in public transport, public transport value of time and public transport improvement preferences using mixed mode data collection. The study concerns Internet and mail-back combinations in the first two cases, and Internet combined with computer-assisted home interviews in the third case. The responses are not directly comparable in the first two cases, as the survey is more customized in the Internet part. In the third case, interviewees are exposed to the same treatment in the Internet case as in the home interview case. For this group, no significant parameter differences were found. Response rates varied between 30 and 43 percent. A little more than half of the

respondents preferred the Internet to the other option. In this case, the sample sizes were 352 people using home interviews and 582 using the Internet.

Sudeshna (2009) points out that several studies find that computer-aided self interviewing techniques are of higher quality than interviewer-oriented methods in several aspects, as a consequence of being subject to much less time pressure when responding to the survey.

As the Internet facilitates customized interviewing in a cost efficient way, it is obvious that using it has a great potential for high quality surveys. Then the issues of low response rates and lack of Internet access need attention, possibly by mixed mode surveys.

The research issues of interest relate, therefore, not only to the performance of different data collection modes, but also to the question of how different data collection modes can be combined. Stopher (2010) maintains that "the travel survey profession needs a controlled experiment in which a test can be made of the effects on survey measurements of using different survey modes, such as mail, Internet, and telephone. The potential increase in response rates from multi-mode surveys also need to be identified better. Theoretically, they should be significant, but this has not been proven as yet."

3 Method

During Oct-Nov 2008 a Swedish national and large-scale Stated Choice survey was conducted for private trips, to be used for inferring the value of travel time (VTT). For this study, different data collection methods were tested in reasonably large pilot surveys before the main study was undertaken. We use data from these pilots and the main survey to study the properties of different data collection methods. We do this by analysing two pilot studies (Internet and telephone) first without parametric constraints and then using parameterized models. We also do this for the main study (using Internet with a telephone non-response follow-up). In the next section, we briefly describe the Stated Choice design and the model specification used for eliciting the VTT distributions. We limit the analysis to the car mode.

3.1 Stated Choice Design and Model Specification

The stated choice experiment described here comprises choices between alternatives differing in two dimensions: travel time and travel cost. The experiments were designed so that the respondent was presented with one reference alternative and one alternative with a change from the reference alternative. The reference trip was the selected trip reported by the respondent. Respondents were also asked to refer to this trip while stating their choices.

To make it possible to capture the well known valuation gap between gains and losses, the design comprises four types of choices, namely 'willingness to pay'- *WTP*-choices and 'willingness to accept' *WTA*-choices, equivalent gain *EG*-choices and 'equivalent loss' *EL*-choices. The *WTP*-choice is presented as a choice between one alternative representing the current trip and one alternative that is faster but more expensive. The *WTA*-type of choice is the exact opposite, including one alternative

representing the current trip and one alternative that is slower but less expensive. The EG type of choice is a choice between one alternative containing the reference trip time and is cheaper, and another alternative that contains the reference cost and is faster. The EL type of choice implies a choice where one alternative has the reference time and is more expensive, and another alternative that has the reference cost and is slower. The four types of choices were presented equally often (two times each in each experiment). In addition, a Contingent Valuation type of question was asked after the last SC question (which was a WTP or a WTA type of question). In that question, the respondent was asked to state the maximum willingness to pay (accept) for the presented time gain (loss). This question was asked in order to find suitable truncation points of the VTT distribution.

The econometric model used in the present analysis is formulated directly in terms of willingness-to-pay and is derived in Fosgerau (2006). The model has proven to fit data well in which respondents make binary choices between trip alternatives that differ only by time and cost. The design is such that in all cases, one alternative is faster but more expensive than the other. The approach can briefly be described in the following way:

Denote the cost of each alternative as c_i and the total travel time as t_i. Denote the individual specific VTT as W. Let y be the choice indicator of each binary choice, defined with the convention that y = 1 if the slow and cheap alternative 1 is chosen, otherwise y = 0. Let $V = (c_1 - c_2)/(t_2 - t_1)$ be the trade-off price of travel time implicit in the choice situation. The experimental design is such that $V>0$ for all observations. We call V the bid. Then an individual will choose the slow alternative if his VTT is smaller than the trade-off price of time (the bid), i.e. if $W<V$. Taking logs and adding an error term leads to the model

$$y = 1\{\log W < \log V + \mu\varepsilon\}. \qquad (1)$$

The error term ε is taken to be iid standard logistic, such that a logit model results. The parameter μ determines the scale. The VTT is parameterized as

$$W = \exp(\beta x + \delta), \qquad (2)$$

where β is a vector of parameters, x is a vector of independent variables and δ is a constant, which may be individual-specific and hence random. This formulation ensures that W is positive, while the ranges of β and δ are unrestricted. The ease with which covariates are incorporated is an important advantage of the present model.

The assumption that W is individual-specific and varies randomly in the population takes care of the correlation of the unobserved heterogeneity arising from repeated observations of the same individuals. The error ε is still assumed to be independently and identically distributed also within individuals.

4 Data Collection

Prior to the main data collection, careful piloting was undertaken with the aim of investigating how two different data collection methods influence the response rate and the VTT. Respondents were interviewed over the Internet or by telephone. Recruitment was made by sampling from the population register. All respondents were asked to report all trips as driver longer than five km in a trip diary. One trip was then randomly selected to serve as the reference trip in the Stated Choice experiment. The lower boundary for the drip distance was set due to the fact that it is not feasible to construct relevant and realistic Stated Choice questions for very short trips. As we were only interested in car trips, we expected a large number of respondents not to be part of the sampling target. For public transport trips (not dealt with here), recruitment was made by intercepting actual trips.

4.1 Pilots

The pilot data collection was undertaken using two different survey methods. Respondents were randomly divided into two groups. For one group, the survey was designed as a telephone survey and for the other group as an Internet survey. Both used exactly the same questionnaire including questions concerning socio-economic information of the respondent and her household and questions related to Stated Choice experiments.

Respondents were recruited by drawing a random sample from the population register. Care was taken to make the telephone and the Internet samples similar in terms of socio-economic background. Telephone numbers was then attached to the telephone sample using a telephone directory including fixed telephone subscriptions and in some cases mobile subscriptions and IP-telephone numbers. For about 30 percent of the sample, the telephone number could not be matched and these were excluded from the sample. The telephone sample consisted of 500 individuals (after telephone number has been attached to the sample) and the Internet sample consisted of 1000 individuals. The larger sample size for Internet was used because of the expectation of low response rate. To guarantee a certain number of responses from elderly individuals, who are said to have less access to Internet, the two groups had an overrepresentation of people older than 50. Also, people below 25 were not included.

Initially, all respondents in the two groups were informed about the survey and invited to participate in a letter sent by regular mail. A trip diary was attached in this letter, for the respondent to fill in their trips on a particular survey day. In the first part of the questionnaire, respondents were asked to report all trips in the survey diary (trips as driver longer than 5 km). One trip was then randomly selected to serve as the reference trip in the Stated Choice experiment.

The Internet group was also given an Internet address to the survey site as well as username and password to login. The telephone respondents instead received sheets where respondents could fill in the time and cost data for the Stated Choice experiments at the interview.

Telephone respondents were then contacted by telephone repeatedly until reached, but at most three times. Finally, 266 respondents were interviewed, resulting

in a response rate of 53 percent. Out if these, 181 respondents had made a valid car trip and completed the SC experiments. Internet respondents not responding after one week were sent a reminder by regular mail. No further effort was made to increase the response rate. One hundred and ninety seven Internet respondents completed the survey and hence the response rate was slightly less than 20 percent. Figure 1 shows how the response rate developed over time, indicating an important effect of the reminder. Out of the 197 respondents, 153 had made a valid car trip and had completed the SC experiments. In addition to the SC experiment related to the car trip, respondents were also asked to complete an experiment related to an alternative mode (bus or train).

The 334 respondents completed 6,081 SC questions, including those for the alternative mode. After cleaning the data, 2,980 observations remained. From these, data collected by Internet consisted of 1,362 responses from 99 individuals. The data collected by telephone consisted of 1,618 responses from 107 individuals. One hundred and twenty eight respondents were discarded from the VTT model estimation because they had reported the reference trip as having too low or too high a speed (making the stated choices inconsistent), or because they did not pay for the trip themselves (making the cost less relevant).

Although the samples were constructed to be similar in terms of socio-economic characteristics, we observe that the final response groups are different in terms of personal characteristics. In the Internet sample, individuals not having access to Internet cannot respond to the survey. In the telephone sample, individuals for whom the telephone number could not be matched with telephone directory are excluded. Hence, different administration methods give rise to differences in the personal characteristics of the respondents, which must be kept in mind when comparing responses between the samples.

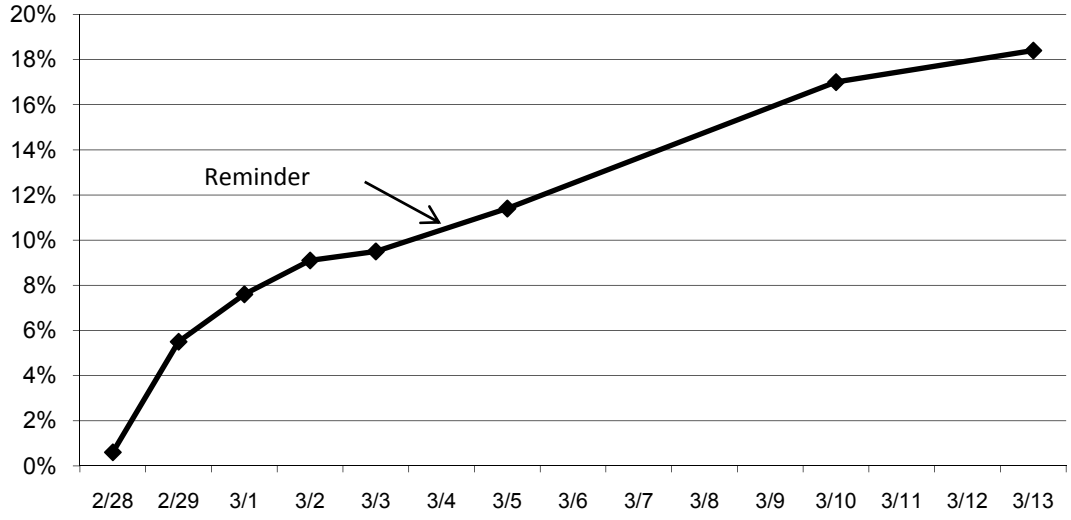

Figure 1: Response rate development. The survey opened in February 28 and reminder was sent out March 4

Note also that the number of reminders is not comparable between the studies administered by Internet and telephone, which are relevant when comparing response rate. Repeated telephone calls are not equivalent to the reminders by mail, since respondents do not know that they are being contacted until they are reached the first time.

4.2 Main survey

From the pilots it was obvious that relying on the Internet would give a very low response rate. But to take as much advantage as possible of the Internet cost efficiency, respondents in the main survey were directed to the Internet as the standard option. To avoid a potential selection bias and low response rate due to sole use of Internet, a full scale non-response survey by telephone complemented the mixed mode survey approach.

As in the pilot surveys, respondents were recruited by drawing a random sample from the population register. Telephone numbers was then attached to all individuals in the sample, resulting in a dropout of 30 percent of respondents for whom the telephone number could not be matched. The final sample comprised 6,000 individuals.

All respondents were initially invited to participate by ordinary mail, in which they were also informed about the survey. The respondents were informed that the survey focused on their driving on a particular survey day. A one-day trip diary was attached in this letter, and respondents were requested to report car trips longer than five km (during which they had been the driver) on the pre-specified survey day.

The respondents were asked to respond to the survey by logging in to the survey site. Respondents were also informed that in the event that they lacked access to Internet they would be interviewed by telephone in a few days. In case there would be a later telephone interview, sheets were attached in the letter, to help presentation of the Stated Choice alternatives. All respondents were compensated with a lottery ticket in advance.

Three days after the specified survey day, non-respondents were contacted by telephone repeatedly until reached, but at most three times. They were requested to respond to the questionnaire over the Internet, but if they stated that they did not want to, or could not (for some reason), respond by Internet, they were given the chance to respond to the survey over the telephone. Respondents still not reached received a reminder by mail with an updated survey day. Two to three weeks later, another wave of telephone reminders was undertaken, similar to the first one, to further increase the response rate.

Out of the sample of 6,000 respondents, 2,040 persons did not fulfil the requirements for participating in the survey since they had not made a car trip longer than 5 km as driver on the pres-pecified survey day. One thousand one hundred and ninety nine persons responded to all SC questions by Internet and 262 persons responded by telephone. This gives a response rate of 59 percent. Since people that had not made any valid car trips most likely felt less obliged to respond than others, this is probably an underestimation of the relevant response rate.

After cleaning the data, 1317 respondents remained, giving 8877 SC observations. 160 respondents were discarded from the VTT model estimation because they had reported the reference trip as having too low or too high a speed, because they did not pay for the trip themselves (making the cost less relevant) or because they had only chosen the left or right alternative through all SC questions (5 telephone respondents and 19 Internet respondents). The reason for discarding respondents with unrealistic travel speeds (implicitly) stated for the reference trip is that the SC experiment then becomes very unrealistic.

As in the pilot studies, the telephone and Internet samples give different results in terms of personal characteristics. The samples are distinguished not so much by Internet access but rather by whether the individuals would actually fill in the Internet survey, which is highly correlated with Internet access.

5 Sample statistics

Personal income is coded into six income intervals. Before-tax income has been set at interval midpoints and then transformed to after-tax income using the tax rates. The sample statistics of the pilot and main survey are summarized in Table 1.

In the pilot, there are only small differences in socio-economic characteristics between the Internet and the telephone groups. Although the travel times do not differ much, distances are shorter in the telephone sample.

In the main survey, the comparison of statistics between the Internet and the telephone samples reveals some significant differences. The respondents in the telephone group have lower income, a lower level of employment and are older. The proportion of men and women is, however, the same in the two groups.

Table 1 Sample statistics for the respondents in the pilot and the main survey

	Woman	Employed	Pensioner	Distance (km)	Time (min)	Income after tax (kSEK/m)	Age	Purpose: Commute	Purpose: Service
Pilot									
Internet	0.44	0.69	0.18	63	105	17	53	0.50	0.18
Telephone	0.39	0.72	0.20	47	97	17	52	0.46	0.17
Main									
Internet	0.58	0.73	0.16	95	81	17	48	0.25	0.48
Telephone	0.59	0.59	0.34	81	73	14	55	0.25	0.45

6 Estimation Results

6.1 Examining the pilot data

Our focus in comparing the two survey methods is the response structures induced by the different survey administration techniques. To explore the VTT distribution revealed by the data, we perform a local constant regression. This method is described in detail in Fosgerau (2007). Following the notation in section 4.1, we may write the individual i's response to the SC questions as $y_i = \mathrm{E(y|V)} + \eta_i$, where η is a random error. By definition, we have that $\mathrm{E}(\eta|\mathrm{V}) = 0$. $\mathrm{E(y|V)}$ equals the probability that the bid is rejected, $\mathrm{P}(\mathrm{W} < V)$, and we may thus think of $\mathrm{E(y|V)}$ as an estimate of the cumulative distribution function of W. We estimate $\mathrm{E(y|V)}$ by regressing on y given V (given some weak conditions). This regression gives us:

$$\mathrm{E(y|V)} = \mathrm{P(y = 1|V)} = \mathrm{P}(\mathrm{W} < V) = \mathrm{F_W(V)}.$$

To clarify, the point, estimate of $\mathrm{F_W(V)}$ is thus constructed by weighting observations in the neighbourhood of the point (the width of the neighbourhood is called bandwidth) using a normal density kernel. The larger the bandwidth we use, the smoother the estimated curve becomes. We use the same bandwidth, 1, for both data sets. The bandwidths are selected to be relatively small so as to preserve the response error, since we are primarily aiming at illuminating the response structures.

Since we have approximately the same number of observations in the two interview modes, we may compare the variances of the responses between the two data sets. Figure 2 demonstrates that response error is larger for telephone interviews. Particularly for high bids, the responses are less homogenous than the responses collected using the Internet, but the response error is larger in the entire range of bids.

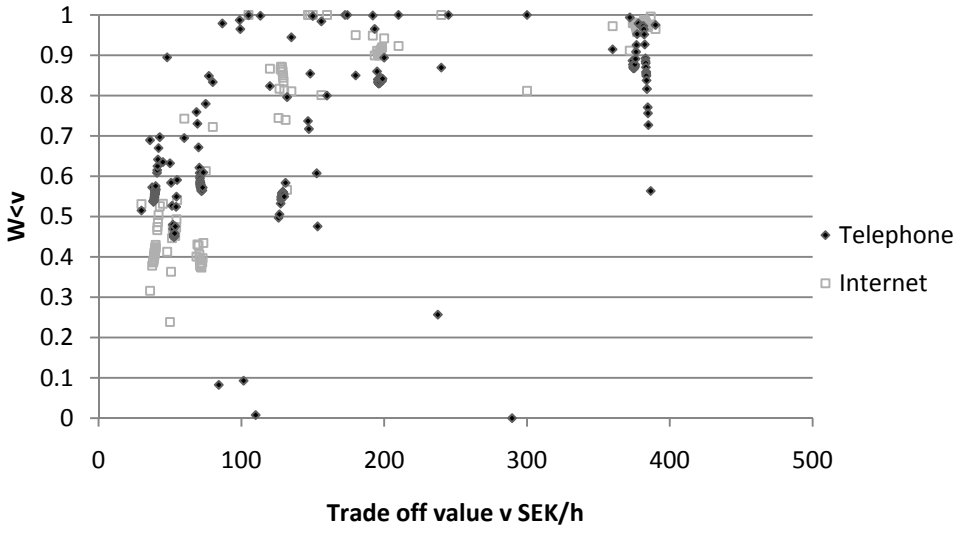

Figure 2 *y* regressed on *v* using a normal density kernel with bandwidth 1

The higher response error in telephone data could be caused by a perceived time pressure when responding over the telephone, where the interviewer is waiting for responses. For this reason, respondents may not take the time to fill in the SC sheet. Another hypothesis is that the response burden becomes larger for telephone respondents, because having the choices read out and filling in sheets induces fatigue in the respondents, which takes focus from the task. These hypotheses, in particular the first one, are strengthened by the fact that Internet respondents use more time to respond to the SC question than telephone respondents, in spite of the fact that telephone respondents also have to take time to listen to the read-out choices and to fill in the SC sheet. Figure 3 shows the response times for all except for the first SC questions for the car mode and the alternative travel mode. (For technical reasons, the response time was not coded for the first choice.)

Interestingly, the response times are consistently lower for the stated choices concerning the alternative mode. This could indicate response fatigue, since these questions are added towards the end of the interview, while the SC concerning the original, car, mode is in the beginning of the interview. Another explanation is that the response times decrease as respondents become more familiar with the task.

The present data only reveal the trade-offs within the range 3 to 390 kr/h. This means that the data cannot determine the tails of the distribution which is a problem if we are interested in determining mean VTT. This problem seems to be relatively small, particularly for Internet respondents since the maximum bid corresponds to approximately the 0.95 quantile according to the figure above. For telephone respondents there is large response variability close to the maximum bid, and the maximum bid therefore corresponds to a smaller quantile of the VOT distribution, approximately the 0.90-0.95 quantile.

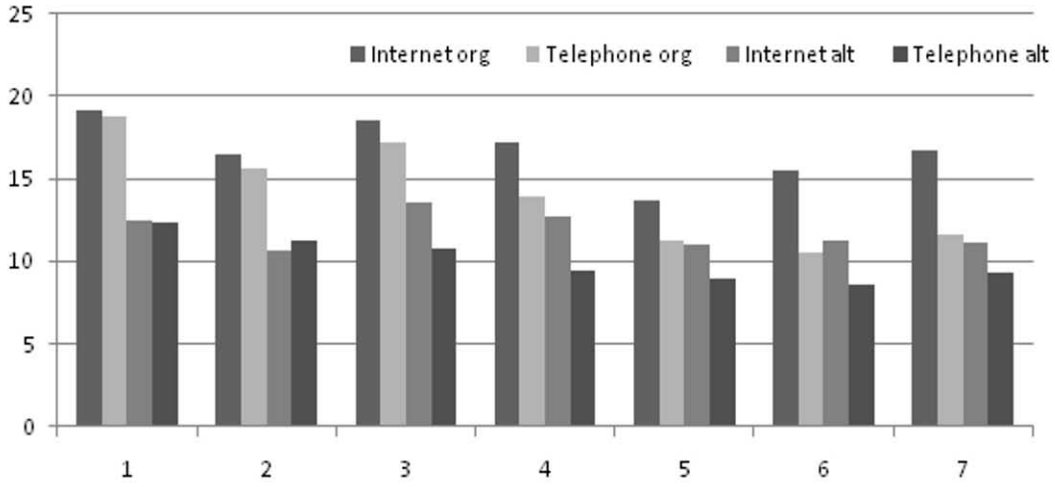

Figure 3 The response times (seconds) for the last seven SC questions for car and the alternative travel mode (the response time was not coded for the first choice)

The number of respondents accepting all bids is small, but somewhat higher for telephone respondents: two percent among telephone respondents and 1 percent among telephone respondents. This is another indicator that we have covered a large part of VTT distribution.

6.2 A Parametric Model Estimation on Pilot Data

Using the econometric model defined in section 3.1, we estimated two models, one of which includes a large number of socio-economic variables in the β-vector (model *1b*). The assumption that x and δ are independent is crucial, implying that the distribution for the VTT is unaffected by a shift in x. Hence, we need to assume that socio-economic covariates are correlated with δ. Model *1a* is identical to model *1b,* but includes no socio-economic variables.

The two models also include dummy variables for EL, EG and WTP-type bids. To capture the fact that VTT increases with the size of the travel time difference (Δt) between the two alternatives in the binary choice, we include the variable log Δt. In particular, it allows for the effect that small travel time changes may have a smaller unit value. The specification also includes the log of the trip distance (*log cost*) and the log of the travel time (*log time*). The δ parameter is taken to follow a normal distribution.

The models are estimated on the joint Internet and telephone samples, but allow for different mean and standard deviations of δ, describing the random variation in VTT, of the two samples. Different scale parameters are also estimated for telephone and Internet responses since the local constant regression plots suggests that there is more response error in the telephone reported responses. The present model specification therefore distinguishes the variance induced by the within-individual error from the variance induced by the variation between individuals, here interpreted as random variation in the VTT.

Table 2 shows the model estimates. The model estimation confirms the findings in the previous section: there is considerably larger error in the data collected by telephone than those collected by Internet. The scale of the error terms is smaller in the responses collected by telephone, implying that the error within individuals is larger. The variance of the value of travel time distribution is also higher for telephone respondents.

Calculating the mean VTT requires an assumption about VTT for the distribution above the maximum bids. Since a larger part of the mass is placed above the maximum bid for telephone respondents, this assumption has a larger impact on the calculated mean VTT for telephone respondents. For this reason, the mean VTT of the telephone and Internet respondents cannot be clearly distinguished. Just to state an example, however, the means of the censored VTT distributions are shown in Table 2. The choice of truncation point, 1500 SEK/h, is guided by responses to the CV question asked in the main survey. To compute the sample means, we average over simulated individual values of time. These calculations indicate no large difference in the mean VTT between Internet and telephone respondents. If the distributions are left uncensored, the VTT are quite close.

Table 2 Parameter estimates for pilot models

Model	**1a**		**1b**	
Number of MLHS draws (Hess et al., 2006):	1500		1500	
Number of observations:	2980		2980	
Number of individuals:	206		206	
Final log-likelihood:	-1091.49		-1075.73	
ρ^2	0.472		0.479	
Adjusted ρ^2	0.465		0.467	
Name	**Value**	***t*-test**	**Value**	***t*-test**
EG, dummy	-0.63	-2.32	-0.63	-2.32
EL dummy	-0.22	-2.37	-0.22	-2.40
WTP dummy	-0.57	-2.55	-0.58	-2.57
Log Δt (minutes)	-0.26	-2.42	-0.27	-2.44
Log cost (SEK)	0.18	1.55	0.16	1.43
Log time (minutes)	0.25	0.95	0.36	1.45
Log Income (SEK/month after tax)			0.05	2.59
No Income coding, dummy			0.94	1.52
Employed, dummy			0.37	1.15
Log Age (years)			-1.02	-2.31
No Age coding, dummy			1.46	0.84
Commuting purpose, dummy			-0.09	-0.30
Recreation purpose, dummy			-0.41	-0.95
School purpose, dummy			-0.45	-0.68
Service purpose, dummy			-0.33	-0.87
Business purpose, dummy			-0.07	-0.12
Important fixed arrival time, dummy			0.09	1.20
Alternative Mode, dummy	0.03	0.38	0.03	0.36
Woman, dummy			-0.17	-0.75
Mean constant (δ) Telephone	-0.53	-2.06	-1.45	-3.22
Mean constant (δ) Internet	0.12	0.56	-0.77	-1.95
Std dev. constant (δ) Telephone	1.69	6.24	1.62	6.13
Std dev. constant (δ) Internet	1.19	6.54	1.03	6.39
Scale Telephone	1.13	0.87	1.12	0.81
Scale Internet	2.19	4.26	2.2	4.28
Mean VTT, truncation point 1500 kr/h(SEK/h), Internet	86		65	
Mean VTT, truncation point 1500 kr/h(SEK/h), Telephone	77		68	
VTT at mean (SEK/h), not truncated, Internet	90		66	
VTT at mean (SEK/h) not truncated, Telephone	93		88	
Standard dev. of VTT at mean (SEK/h) Internet	168		108	
Standard dev. of VTT at mean (SEK/h) Telephone	372		289	

The mean VTT values differ somewhat between model 1a and model 1b. Presumably, the difference depends on misspecification of the model when including covariates. As an example, age and income are restricted to having a constant elasticity of VTT. Also, the model assumes that the covariates are independent of the VTT distribution, which might also be too strong an assumption for the model including socio-economic variables. Consequently, we believe more in the VTT computed with model 1a. Model 1b may state an example of the risk of including many socio-economic covariates in the model specification.

6.3 Examining the main survey data

Figure 4 compares the response structure of the data in the same fashion as for the pilots, performing a local constant regression. We still use the same bandwidths for telephone and Internet respondents.

The figure shows quite clearly that the VTT distribution for telephone respondents is more skewed to the right, that is, a larger part of the mass is located at VTT less than 50 SEK/h. Hence, we expect that the VTT is lower for the respondents interviewed by telephone. The figure also indicates larger response variability, and hence a larger response error, for telephone respondents than Internet respondents, just as in the pilots.

Only 15 telephone and Internet respondents out of 1,317 accepted all bids, so the coverage of the distribution is quite satisfactory.

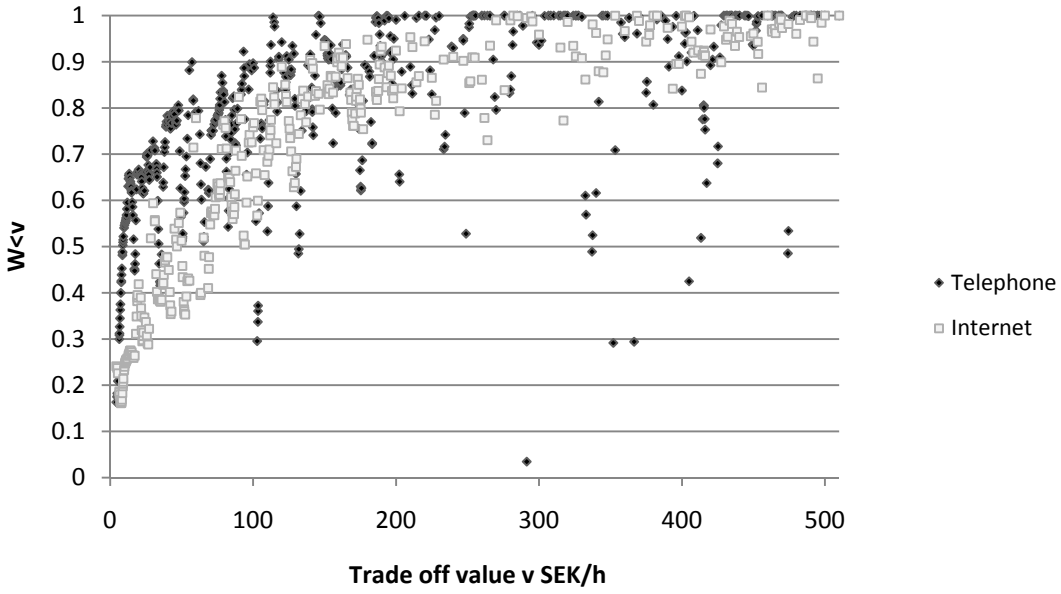

Figure 4 *y* regressed on *v* using a normal density kernel with bandwidth 1

6.4 Parametric Model estimation main data

The parametric models are defined in the same way as for the pilots. In Table 3 we present the estimated parameters of two model variants, one without and one with socio-economic covariates. We also present the mean VTT truncated and not truncated.

As in the pilot, the mean VTT for the main study was computed by simulation from individual observations. The mean VTT is approximately 45 percent higher for Internet respondents, and the standard deviation of the VTT is smaller for Internet respondents. In the pilot, the difference between Internet and telephone in terms of mean VTT was small, but here the difference is apparently larger. To further illustrate the distributional differences of the two subsamples, the probability density function is plotted in Figure 5. There seems to be a mass point close to the lowest bids for telephone respondents in the choice experiment. This is also the case for Internet respondents, but it is less outspoken.

One reason for the VTT to be lower for telephone respondents is that refusal to respond by Internet may be associated with covariates which are positively correlated with VTT. Refusal to respond by Internet could also be correlated with non-observed variations in VTT. A key point of interest is if the differences in VTT can be explained by the difference in socio-economic composition of the samples, implying that the unobserved difference in VTT is less important. In this case, information about the distribution of the socio-economic variables for non-respondents could be used to achieve an unbiased estimate of the VTT for the total population just by reweighting the sample with socio-economic distribution of the full sample.

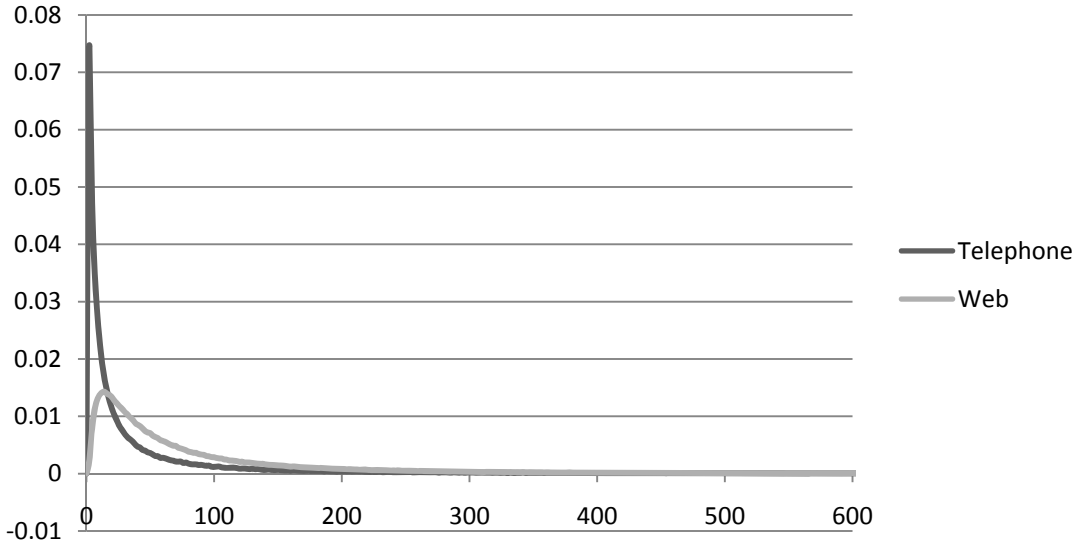

Figure 5: Probability density function for Internet and telephone subsamples

Table 3 Parameter estimates for the main survey

Model	**2a**		**2b**	
Number of MLHS draws (Hess et al., 2006):	1500		1500	
Number of observations:	8877		8877	
Number of individuals:	1317		1317	
Final log-likelihood:	-3225.89		-3173.18	
ρ^2	0.476		0.484	
Adjusted ρ^2	0.474		0.481	
Name	**Value**	***t*-test**	**Value**	***t*-test**
EG, dummy	-0.35	-5.73	-0.35	-5.70
EL dummy	-0.18	-3.01	-0.18	-2.94
WTP dummy	-0.49	-7.76	-0.48	-7.73
Log Δt (minutes)	0.16	2.15	0.15	2.10
Log cost (SEK)	0.05	0.88	0.027	0.45
Log time (minutes)	0.18	1.18	0.31	2.16
Log Income (SEK/month after tax)			0.46	5.03
No Income coding, dummy			-0.11	-0.56
Employed, dummy			-0.14	-1.27
Log Age (years)			-0.73	-5.35
No Age coding, dummy			-2.06	-3.35
Commuting purpose, dummy			0.20	1.23
Recreation purpose, dummy			-0.14	-0.85
School purpose, dummy			0.24	0.44
Service purpose, dummy			-0.17	-1.08
Stockholm resident, dummy			0.27	2.39
Children in household, dummy			0.11	1.14
Mean constant (δ) Telephone	-0.88	-3.99	1.46	3.24
Mean constant (δ) Internet	0.30	2.94	2.39	6.00
Std dev. constant (δ) Telephone	1.80	10.5	1.67	10.4
Std dev. constant (δ) Internet	1.12	25.6	1.05	24.9
Scale Telephone	1.08	0.90	1.08	0.96
Scale Internet	1.82	14.5	1.82	14.4
Mean VTT, truncation point 1500 kr/h(SEK/h), Internet	93		109	
Mean VTT, truncation point 1500 kr/h(SEK/h), Telephone	65		71	
VTT at mean (SEK/h), not truncated, Internet	94		111	
VTT at mean (SEK/h) not truncated, Telephone	73		81	
Standard dev. of VTT at mean (SEK/h) Internet	138		120	
Standard dev. of VTT at mean (SEK/h) Telephone	336		234	

To get some idea of the extent to which the differences in socio-economic status explain the difference in VTT between the samples, we investigate the extent to which the model parameters estimated for the Internet sample can predict the VTT for telephone respondents.

Predicting the mean VTT of the telephone sample by applying the parameters of model 2b (including all covariates) estimated on the Internet sample gives a mean VTT of 89 SEK/h. Applying the parameters estimated on the telephone sample of model 2b to the same sample gives a mean VTT of 81 SEK/h (see Table 3). Hence, the Internet model only overestimates the mean VTT of the telephone sample by 9 percent. Given that the mean VTT is 45 percent higher in the Internet sample than the telephone sample, we can conclude that most of this difference can be explained by socio-economic differences.

A relevant question is then if the extra effort to perform full telephone interviews is worth the costs. The above result suggests not. To avoid including many covariates in the model, the models could use a reweighted sample that matches the socio-economic distribution of the full sample.

7 Conclusions

We will here discuss the three key issues stated in the introduction. First, does the error in responses to SC questions answered by telephone and by Internet differ? According to the pilot, including one survey administered by Internet and one by telephone, there is a considerably lower response scale in the telephone sample. This difference could, in theory, be due to differences in the personal characteristics of the samples caused by the differences in the administration method, rather than the form of administration itself. However, this does not seem to be very likely, given that the socio-economic characteristics of the two samples are similar and that the response times are consistently shorter for telephone respondents. Also, we make the same observation in the main study, in which the personal characteristics between Internet and telephone respondents are indeed different, but in a different way than in the pilot. This suggests that the time pressure caused by the interview situation prevents respondents from taking the time to fill in the SC sheet in order to see both alternatives, and therefore not considering the alternative enough.

Second, how is the response rate affected when using Internet rather than telephone surveys? Response rate drop considerably when Internet is used solely, but much less so if telephone reminders are used. Part of the drop is due to the fact that Internet may not be available to a certain percentage of the Internet sample and they cannot respond for this reason. However, lack of Internet access could not be the only reason for the drop in response rate in the present surveys, given that Internet is available to 82 percent of Swedes in the relevant age group.

Third, do respondents declining to respond to survey questions by Internet but not by telephone, differ from others in terms of socio-economic characteristics or VTT? To respond to this question, we turn to the main study. The respondents declining to respond by Internet, but not by telephone, have a different socio economic status, characterized by lower incomes and level of employment. They have also lower

average VTT, which to a very large extent are explained by the difference in socio-economic status. We stress that Internet and telephone respondents in the mail survey are distinguished by whether they would fill in the Internet survey or not. This does not necessarily coincide with Internet access although it would obviously be highly correlated with it. The reason for non-response by Internet is less relevant for the researcher. We are ultimately interested in how different survey administration techniques affect responses, either because of the administration methods themselves or because of the differences in sample characteristics these give rise to.

There are no significant differences between telephone and Internet respondents' VTT in the pilot survey, because the response frequency in the Internet sample is much lower than the Internet access in the target population. The non-responses in the Internet sample for which telephone reminders are not used are therefore due to various factors and not only to lack of Internet access. The non-responses in the Internet sample in the main survey are, however, to a much larger extent due to lack of Internet access.

The analysis demonstrates further that there is more variation in the VTT within the telephone sample, in other words a more skewed distribution of the VTT. This is the case for the main survey as well as for the pilots. The results suggest that Internet respondents constitute a more homogeneous group than telephone respondents, seemingly because of selection effects.

To conclude this paper, the relevant question is then if the extra effort to perform full scale follow-up interviews by telephone is worth the costs. Our findings suggest that we can do reasonably well without full telephone follow-up interviews as long as information of the socio-economic characteristics of non-respondents is available. However, it appears that a high Internet accessibility in the target population is not enough but also personal reminders are important to achieve satisfactory response rates. Our results suggest that socio-economic characteristics can then be used to predict the VTT of the Internet survey non-respondents. To control for the response bias cased by administering the survey by Internet, a reweighted sample that matches the socio-economic distribution of the full sample should be used in the estimation.

We stress that the importance of follow-ups and options to respond by telephone might be situation-specific and dependent on the extent of Internet access in the target population. Also, the fact that respondents were knowingly subject to full telephone interviews if they declined to partake in the Internet survey, might have had a larger impact on response rate than if only using telephone calls as reminders.

Apart from providing a cheap way of administering SC customized interviews, the Internet is also preferable to the telephone in terms on quality of responses, since reading out the choices over the telephone, instead of delivering them directly on the computer screen, seems to induce more response error.

The most common approaches to data collection seem to be to send the material to the respondent after a scoping interview. The material can be sent electronically or by mail, depending on the need of flexibility in the survey questions. If undertaken by telephone, the scoping interview then has the same function as the important personal reminders in the present survey, and thus resembles the method that we suggest is the

most efficient method. However, it saves some telephone calls to use reminders instead of scoping interviews.

References

Fosgerau, M., 2006. Investigating the distribution of the value of travel time savings. Transportation Research Part B, 40(8), 688–707.

Fosgerau, M. and M. Bierlaire, 2007. A practical test for the choice of mixing distribution in discrete choice models. Transportation Research Part B, 41(7), 784–794.

Hess, S., K. E. Train, and J. W. Polak, 2006. On the Use of a Modified Latin Hypercube Sampling (MLHS) Method in the Estimation of a Mixed Logit Model for Vehicle Choice. Transportation Research Part B, 40(2), 147-163.

Nossum. Å., 2005. Stated Preference Surveys on Internet – an Effective Method for Finding Passengers' Preferences? *TOI Report* 723/2005.

Olsen, S. B., 2009. Choosing Between Internet and Mail Survey Modes for Choice, Experiment Surveys Considering Non-Market Goods, S*pringer.*

Sethuramana, R., R. A., Kerin and W. L., Cron, 2005. A field study comparing online and offline data collection methods for identifying product attribute preferences using conjoint analysis. *Journal of Business Research* 58 602– 610.

Stopher, P., 2010. The Travel Survey Toolkit: Where To From Here? in P. Bonnel, M. Lee-Gosselin, J. P. Zmud and J.-L., Madre, Transport Survey Methods: Keeping Up with a Changing World , Emerald Group Publishing Limited, United Kingdom, 15-46.

Sudeshna, S. 2009. Data collection Techniques – Past, Present and Future. Resource paper presented at the 12th IATBR conference in Jaipur, 2009.